

AI Weekly: YC QM, Vids Avatars, Coding-Tool Consolidation, Mem0, Azure \$100B, ARC Harness Tricks, Superlogical, DeepSWE, MCP Stateless, Buzz BYOH, Open Weights, and X Growth Radar

Weekly Digest · August 1, 2026 · 24 min · July 28, 2026 – August 1, 2026

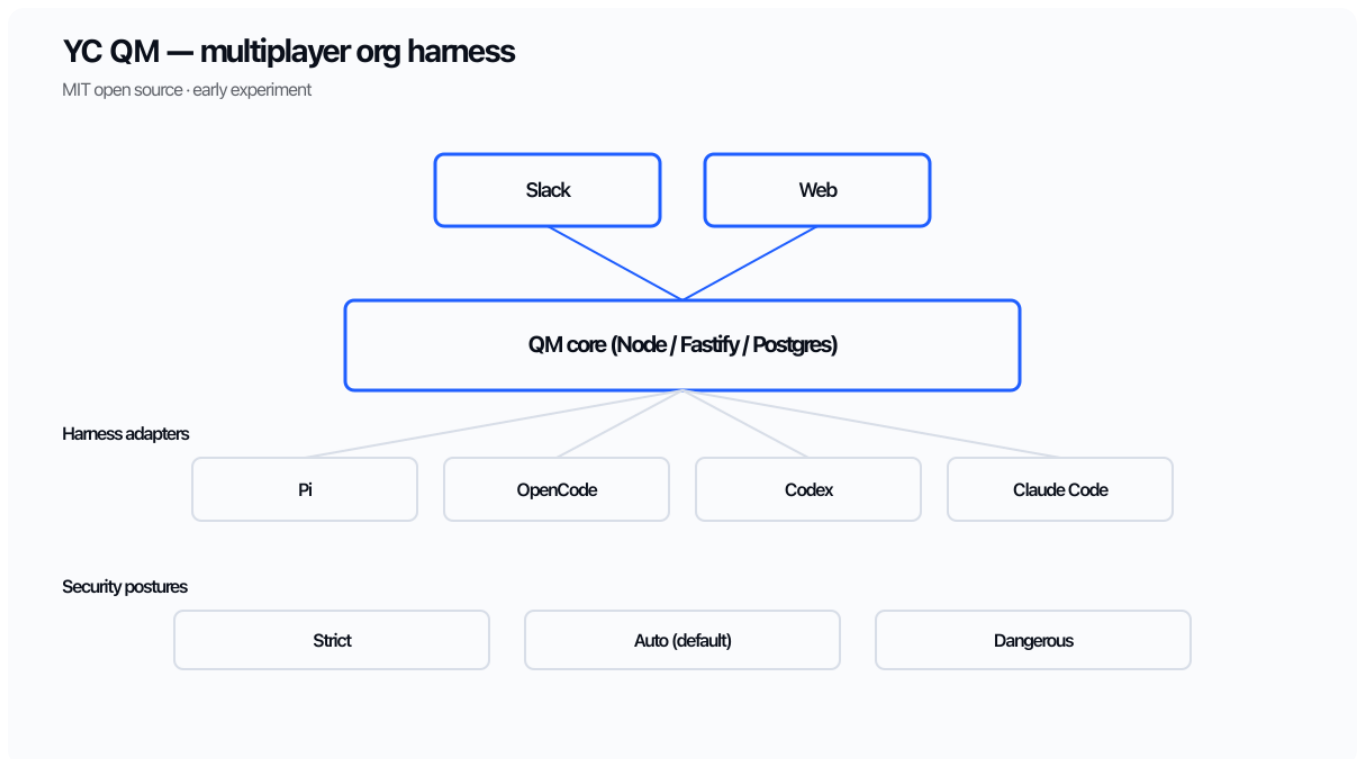
Between 28 July and 1 August 2026, twelve AI stories stood out — and they are easier to follow if you know what each one is really about. First, Y Combinator open-sourced QM, a multiplayer agent harness for startups. Second, Google Vids shipped Gemini Omni video generate/edit plus personal avatars. Third, a viral “end of an era” coding-tools list needed fact-checking against primary deals. Fourth, Mem0 published a first-party Claude Code memory demo. Fifth, Microsoft reported FY26 results with Azure past \$100B. Sixth, OpenAI showed how two Responses API settings change GPT-5.6 Sol’s ARC-AGI-3 public scores. Seventh, Hashimoto and co-founders announced Superlogical — still waitlist-only. Eighth, DeepSWE posted long-horizon coding numbers for Opus 5 and Sol. Ninth, MCP 2026-07-28 made the protocol’s core stateless, with Claude rolling product support. Tenth, Buzz Desktop v0.5.0 added Bring Your Own Harness via ACP. Eleventh, Anthropic restated its open-weights policy stance. Twelfth, an X follower-growth ranking is useful only as an attention radar. Below, each section explains the idea in plain language, then what happened, then why it matters. Sources are at the bottom.

1. YC open-sources QM: a multiplayer agent harness for startups

Most coding agents behave like personal assistants: one human, one chat, one sandbox. An org harness is different — it treats agents as shared company infrastructure, with per-person and per-room memory, permissions, files, and a durable place to run work.

Y Combinator open-sourced QM (short for quartermaster) under MIT in late July 2026 (yc-software/qm around 29 July; social coverage near 31 July). Surfaces are Slack and web with the same identity and config. Harness adapters in the README include Pi,  Code, Codex, and Claude Code driving one core. The stack is TypeScript on

Node/Fastify with Postgres; operators deploy into their own Fly or AWS account. Org-chosen security postures — Strict, Auto (default), and Dangerous — sit on top of a predeclared command policy that always applies. YC's own framing: early, buggy, and dogfooded across accounting, legal, events, and engineering — including building QM itself.



QM overview: shared Slack + web identity, pluggable harness adapters, and org-chosen Strict / Auto / Dangerous postures. See Sources below. [Source](#)

Why it matters: YC is publishing the org-scoped harness it actually uses, not another personal chatbot. Read it as infrastructure signal and early open source — not a finished SaaS.

QM = multiplayer org harness (Slack + web), MIT, model/harness-agnostic adapters.

Postures: Strict / Auto / Dangerous; command policy always on.

YC: early + buggy — useful to study, not a production warranty.

2. Google Vids: Gemini Omni video + personal avatars

Google Vids has been a Workspace video tool. Gemini Omni pushes it toward prompt-driven generate and edit; personal avatars add account-bound likeness so someone can deliver a scripted message without sitting on camera.

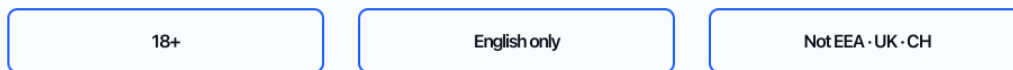


On 16 July 2026, Google announced Gemini Omni in Vids: generate video from natural-language prompts (with optional image references) and chat-style step-by-step edits on Omni-generated or phone-shot clips. Personal avatars capture the account holder's face and voice through a secure Google Account process; scripts drive the avatar. Workspace materials name the model Gemini Omni Flash. Eligibility covers Google AI Pro and Ultra plus listed Workspace SKUs, on by default for eligible customers with admin controls. Avatar gates: 18+, English only, and not available in the EEA, Switzerland, or the United Kingdom. Every clip carries SynthID. Rollout: Rapid Release from 16 July 2026; Scheduled Release from 5 August 2026.

Google Vids — Gemini Omni + personal avatars



Avatar gates



Vids + Gemini Omni: prompt generate/edit, then account-bound personal avatars with SynthID on every clip. See Sources below. [Source](#)

Why it matters: Workspace video is competing with specialist avatar tools from inside Google's compliance envelope — age, region, and watermarking are part of the product, not footnotes.

Omni = prompt video generate/edit; avatars = account-holder likeness only.

Avatars: 18+ / English / not EEA·UK·CH; SynthID on clips.

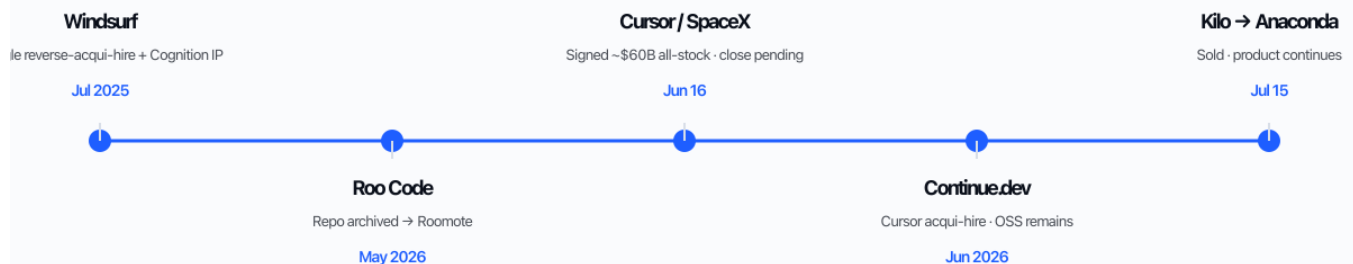
Rapid Jul 16; Scheduled Aug 5 — Workspace Updates is the admin-precise source.

3. ⚡ ding-tool consolidation: fact-check the “end of an era” list

AI coding tools are consolidating — reverse-acqui-hires, pivots, and mega-mergers. Social lists are useful hooks and terrible sources of truth. This section checks the common Jul 30 list items against primary reporting.

Windsurf (Jul 2025): two-step outcome, not a single “sold” line — Google reverse-acqui-hired leadership/R&D with a non-exclusive tech license; Cognition then acquired the remaining IP, product, brand, business, and non-Google employees. **Roo Code (May 2026):** extension repo archived ~15 May; team pivoted to Roomote. **Cursor / Anysphere (Jun 2026):** on 16 June, SpaceX and Anysphere signed a merger at an implied ~\$60B all-stock equity value — expected close Q3 2026 subject to approvals; signed, not confirmed closed. **Continue.dev (Jun 2026):** Cursor acqui-hire shape; standalone product wound down, OSS remains. **Kilo Code (Jul 2026):** Anaconda acquired Kilo on 15 July 2026; products, plans, and support continue for now — sold, not dead. **Cline:** still operating; OpenAI hired ≥7 staffers for Codex (Jan 2026) with no formal acquisition.

Coding-tool consolidation (fact-checked)



Consolidation timeline: Windsurf split (2025), Roo→Roomote, Cursor–SpaceX signed ~\$60B (close pending), Continue into Cursor, Kilo→Anaconda on Jul 15, Cline still independent. See Sources below. [Source](#)

Why it matters: the coding IDE/agent layer is concentrating, but tweet shorthand erases deal structure. Windsurf was a split; Cursor is a signed mega-merger pending close; Kilo sold into Anaconda while shipping; Cline is still a company.

 Hook lists ≠ primary sources — verify each row.

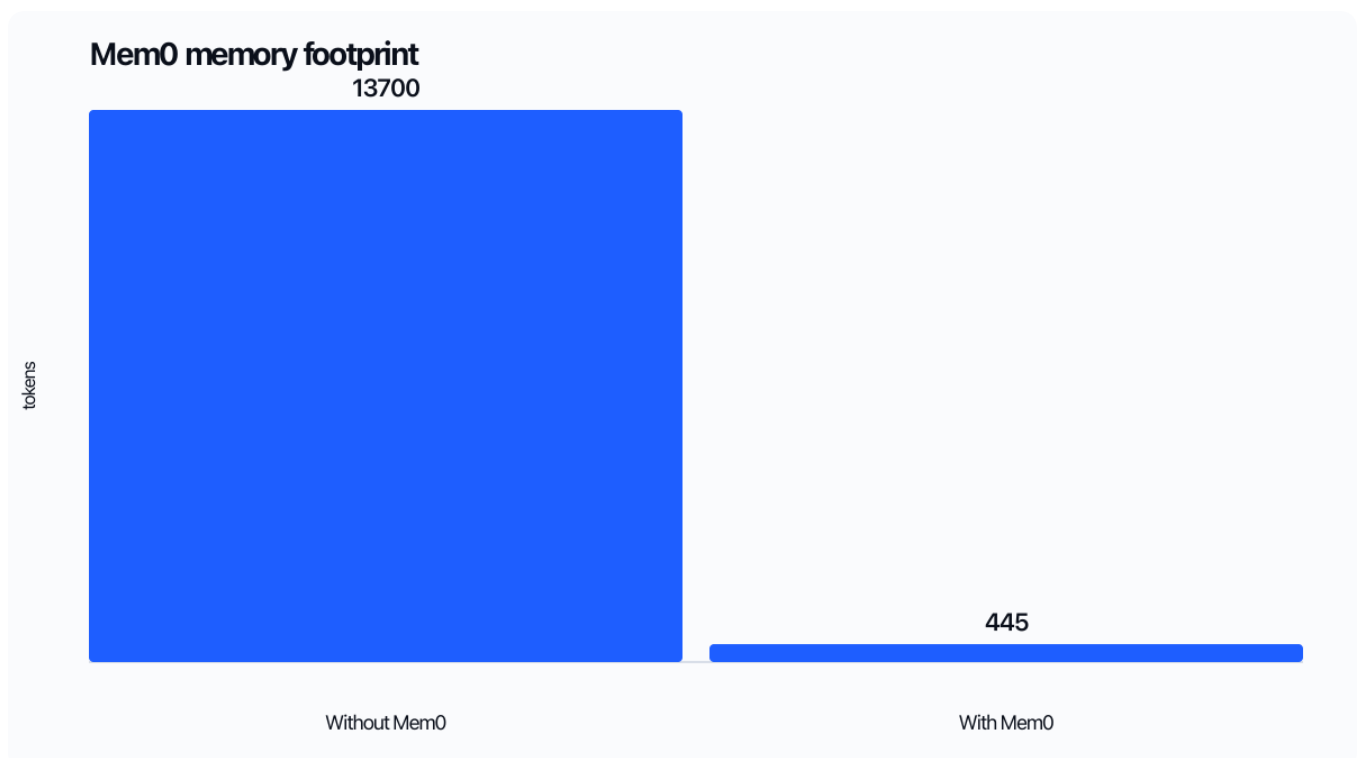
Cursor: signed ~\$60B SpaceX merger (Jun 16), close not confirmed here.

Kilo: sold to Anaconda Jul 15, product continues; Cline: ≥ 7 to OpenAI, no formal acquisition.

4. Mem0: ~97% smaller memory footprint — vendor demo, read carefully

Coding agents often load memory files into every turn. That burns context even when most of the file is irrelevant. A retrieval memory layer tries to pull only the chunk that matches the current prompt.

Mem0 published a first-party side-by-side on Claude Code (v2.1.209) with its marketplace plugin (lifecycle hooks + MCP): same repo, same prompt, with versus without Mem0 (26 existing memories). Measured via Claude Code's `/context`, the memory slice shrank sharply when unconditional file loads were replaced by on-demand retrieval of a relevant chunk. A second experiment claimed preferences stored in Mem0 surviving `/clear`, while the no-Mem0 path forgot them. Mem0's own FAQ frames this as a real-repo Claude Code test — not their published LoCoMo/LongMemEval numbers — and warns against extrapolating from 26 memories to thousands.

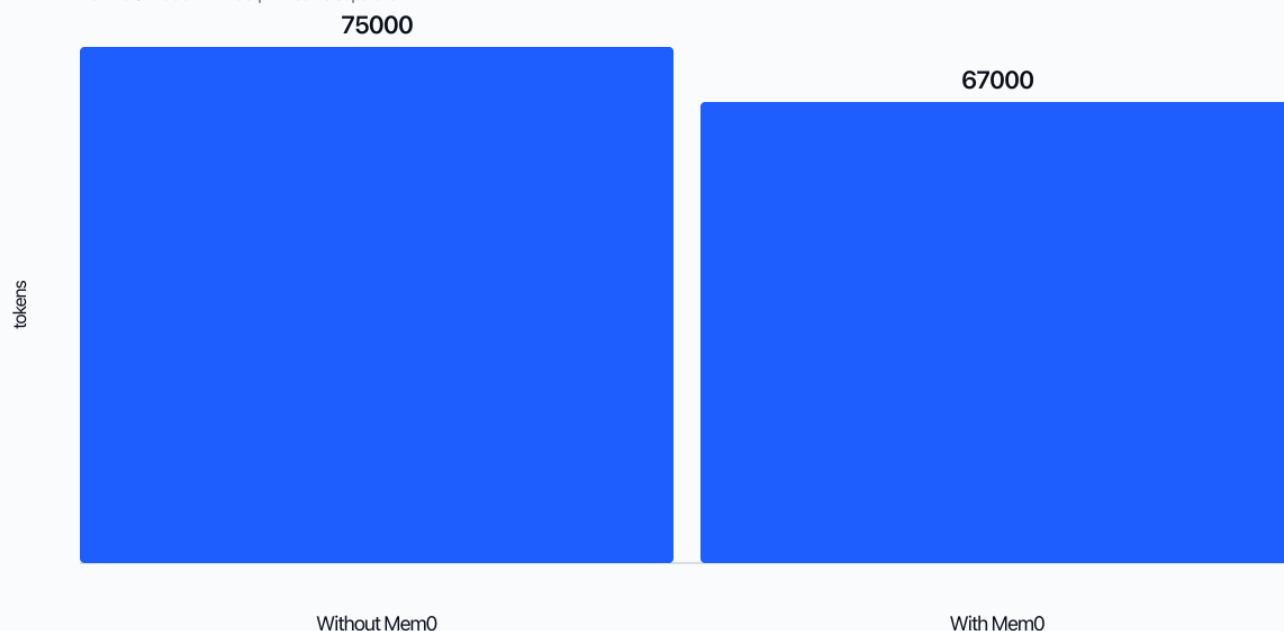


Memory footprint only: ~13,700 → ~445 tokens (~97% smaller) in Mem0's first-party Claude Code demo. This is the 97% claim — not total session tokens. See Sources below. [Source](#)



Mem0 total session tokens

not the 97% claim — footprint cut is separate



Total session tokens in the same table: about 75k → 67k — a much smaller drop. Do not read this chart as the 97% claim (that figure is memory footprint only). See Sources below. [Source](#)

Why it matters: the useful lesson is footprint versus total tokens. The 97% headline is real for the memory slice in their demo — and easy to misread as “97% cheaper Claude Code.” Label it first-party marketing.

~97% = memory footprint (13.7k → 445), not total session tokens (~75k → 67k).

Vendor study on Mem0’s own plugin — not an independent audit.

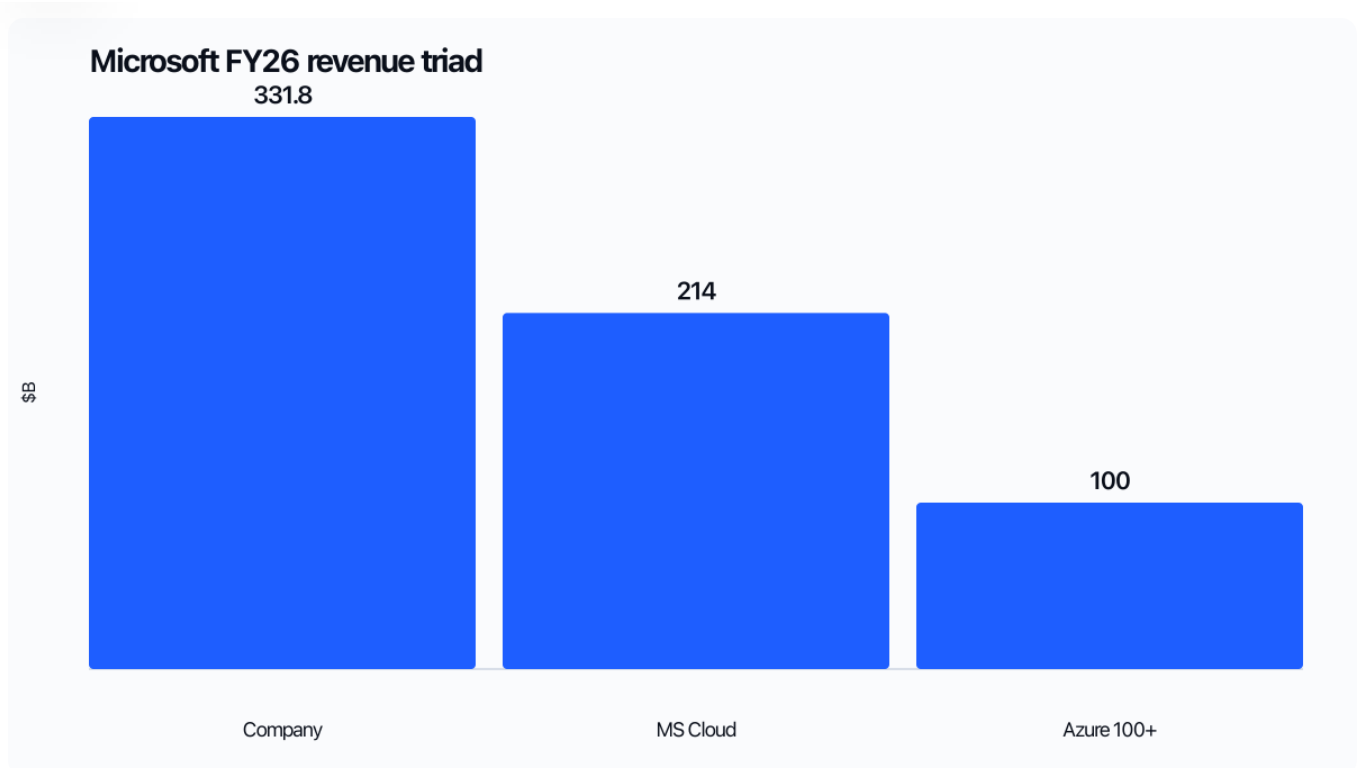
Takeaway: retrieve-what’s-needed vs always-load files.

5. Microsoft FY26: Azure past \$100B, Copilot past 30M seats

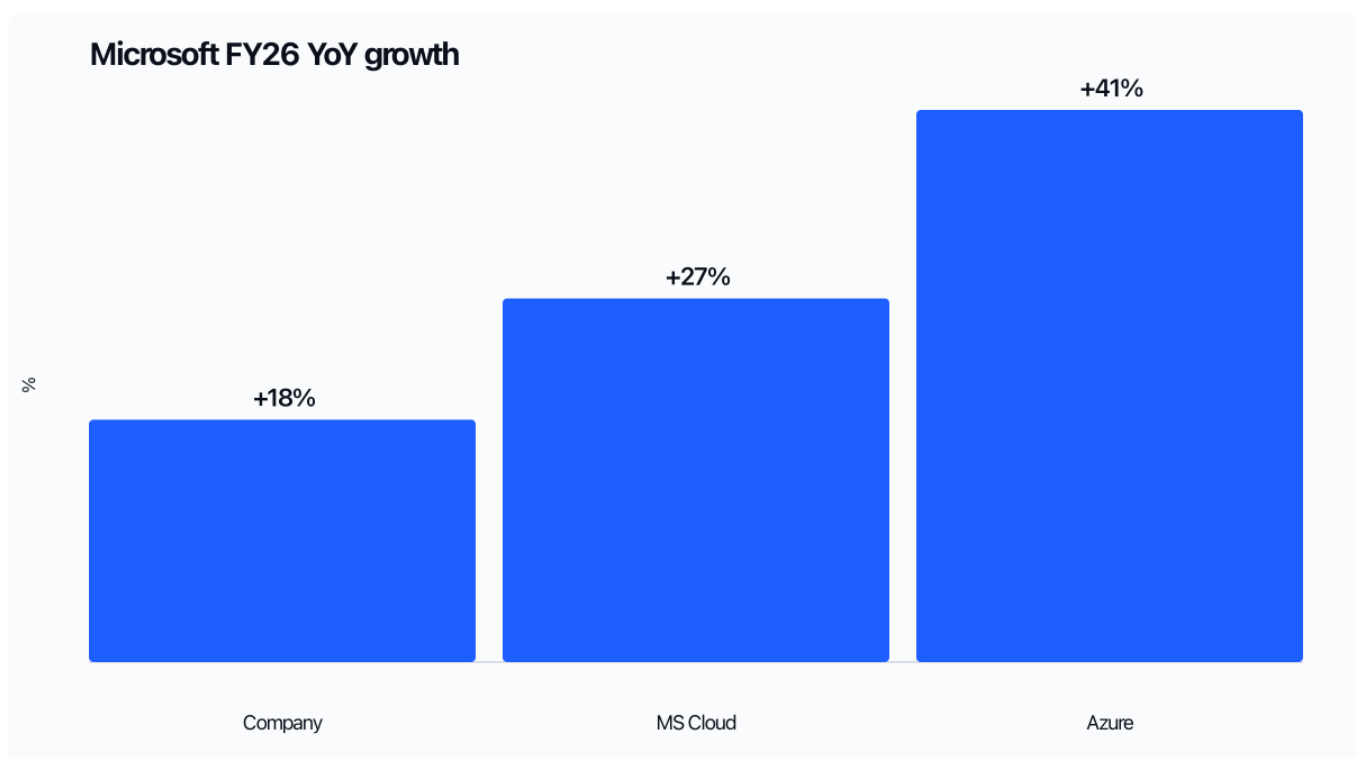
Cloud and Copilot numbers are the commercial backdrop for the week’s agent and video stories. They measure demand and scale — not a stock tip.

On 29 July 2026, Microsoft reported results for the quarter and fiscal year ended 30 June 2026. Company FY26 revenue was **\$331.8B**, **+18%** year over year. Microsoft Cloud surpassed about **\$214B** for the year, **+27%**. Nadella said Azure revenue surpassed **\$100B** for the first time in the fiscal year, **+41%** for the year. Q4 “Azure and other cloud services” grew **+43%** — not a contradiction with the FY +41% figure; Q4 grew faster than the blended full-year rate. Microsoft 365 Copilot reached **over 30 million** paid seats.





FY26 triad: company revenue \$331.8B, Microsoft Cloud ~\$214B, Azure past \$100B. Educational corporate results — not investment advice. See Sources below. [Source](#)



FY26 growth: company +18%, Microsoft Cloud +27%, Azure +41% (Q4 Azure line +43%). See Sources below. [Source](#)

Why it matters: infrastructure CapEx and seat counts explain why harnesses, MCP, and Workspace video keep shipping. This section is educational corporate reporting only — not investment advice.

FY  revenue \$331.8B (+18%); MS Cloud ~\$214B (+27%).

Azure: surpassed \$100B FY (+41%); Q4 Azure line +43%.

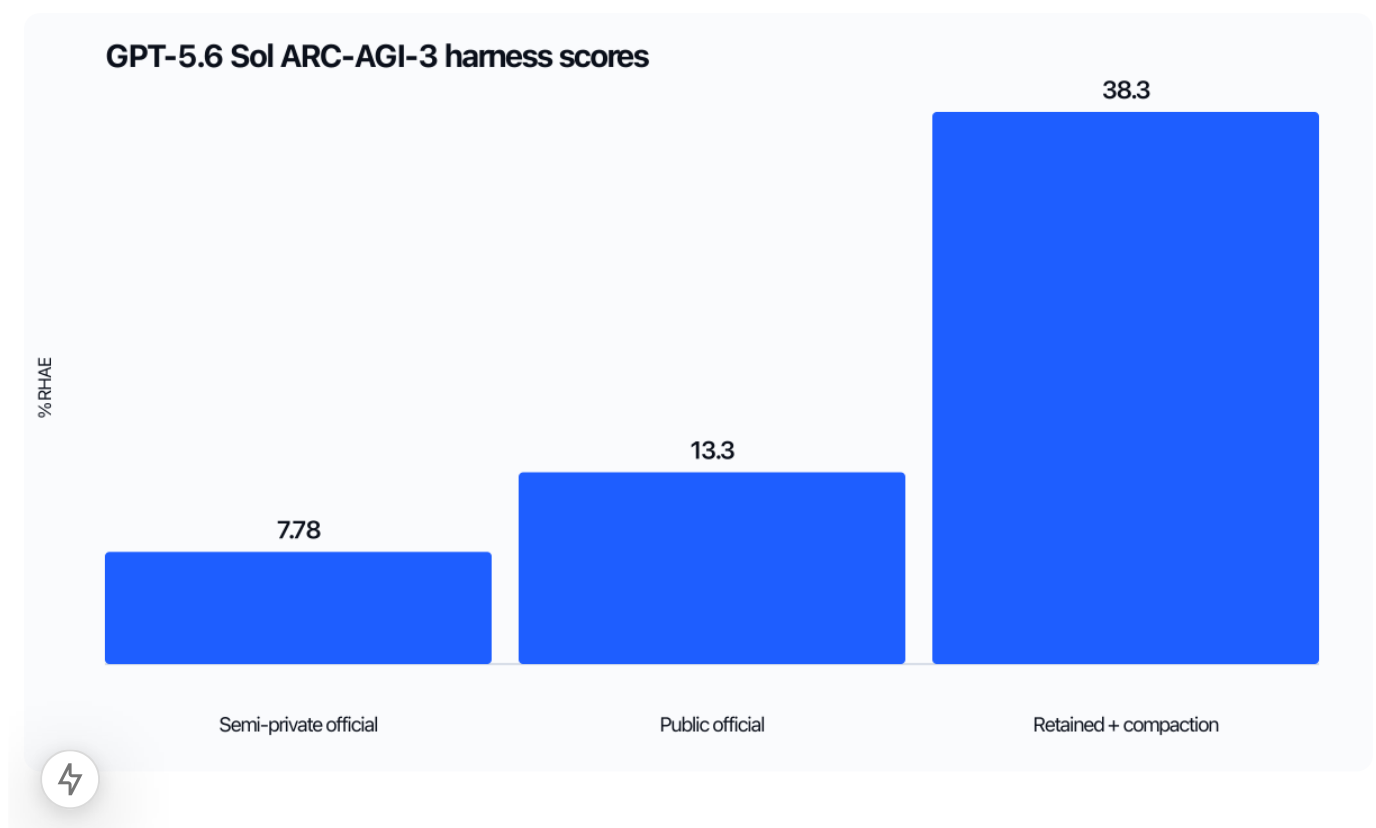
Copilot >30M paid seats — not a buy/sell recommendation.

6. GPT-5.6 Sol on ARC-AGI-3: harness settings change the score

Benchmarks measure models and harnesses together. ARC-AGI-3 scores are Relative Human Action Efficiency (RHAE). Changing how reasoning and context survive across turns can move the number without changing weights.

OpenAI's 29 July 2026 post reports GPT-5.6 Sol at **13.3%** on the official public harness versus **38.3%** on the public set with two Responses API settings: retained reasoning across turns, and compaction instead of rolling truncation. OpenAI also reports roughly **6× fewer** output tokens with those settings. Official ARC Prize pages put Sol Max around **13.33% public** and **7.78% semi-private** — the ~7.8% figure people quote is the semi-private class, not the 38.3% harness run.

François Chollet's reported framing: custom benchmark-only harnesses or benchmark-format knowledge are off-limits; **general-purpose API settings** available to all users can be acceptable if settings and cost are disclosed. Do not mix official, semi-private, and OpenAI self-reported harness scores into one league table.



Three separate Sol ARC-AGI-3 numbers — do not blend: official semi-private ~7.78%, official public ~13.3%, Responses API (retained reasoning + compaction) ~38.3% on the public set. See Sources below. [Source](#)

Why it matters: “Sol tripled ARC” is incomplete without naming which harness. Production-realistic API settings and provider-neutral official harnesses answer different questions.

Official public ~13.3%; retained reasoning + compaction ~38.3% (OpenAI, public set).

Semi-private official ~7.78% — different number, do not blend.

Chollet: general-purpose disclosed settings OK; still not the official harness.

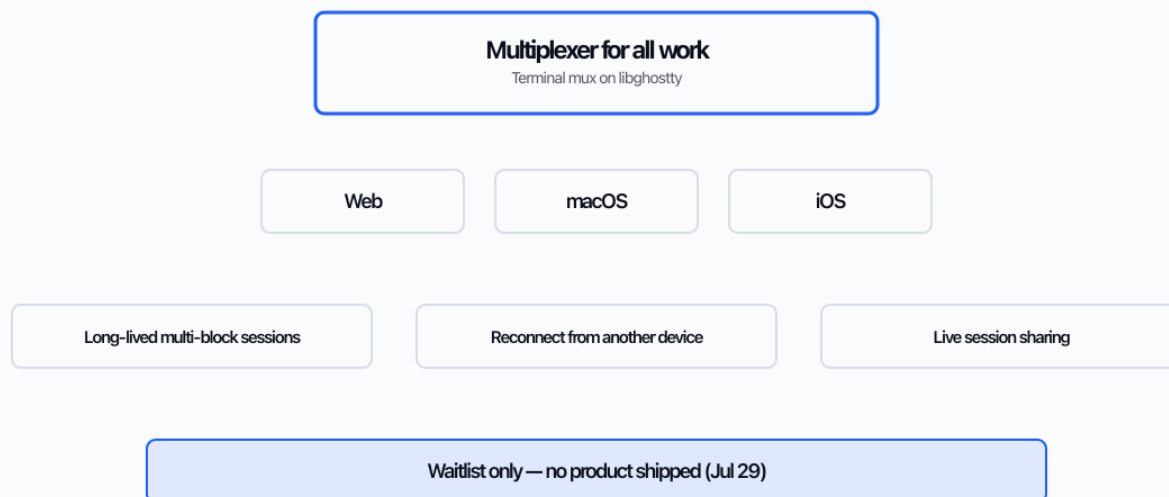
7. Superlogical: a multiplexer for all work — waitlist only

Developers already juggle terminals, agents, CI, and production dashboards. A “multiplexer for all work” is a durable session layer meant to span interactive, automatic, and production work — instead of a pile of disconnected tools.

On 29 July 2026, Mitchell Hashimoto, Jack Pearkes, Alasdair Monk, and Hector Simpson announced Superlogical. The first product is a **terminal multiplexer**: long-lived multi-block sessions, reconnect from another device, web plus native macOS/iOS, live session sharing, built on **libghostty**. Ghostty itself remains a non-profit; Superlogical will upstream shared terminal work. Shipping status as of that announcement: **no product shipped** — beta waitlist only. Funding names circulate (Notable Capital, Amplify Partners, and several angels); amount and valuation are not specified on the primary pages we cite.



Superlogical — session multiplexer



Superlogical thesis: a multiplexer for all work — first ship target is a terminal mux on libghostty. Waitlist only as of Jul 29; not a shipped product. See Sources below. [Source](#)

Why it matters: session durability for humans and agents is becoming its own product category next to QM and Buzz. Superlogical is a high-profile team and thesis — not a downloadable tool yet.

Thesis: multiplexer for all work; first ship target = terminal mux on libghostty.

Founders: Hashimoto / Pearkes / Monk / Simpson.

Waitlist / hiring as of Jul 29 — do not treat as shipped.

8. DeepSWE: Opus 5 and Sol on contamination-free long-horizon coding

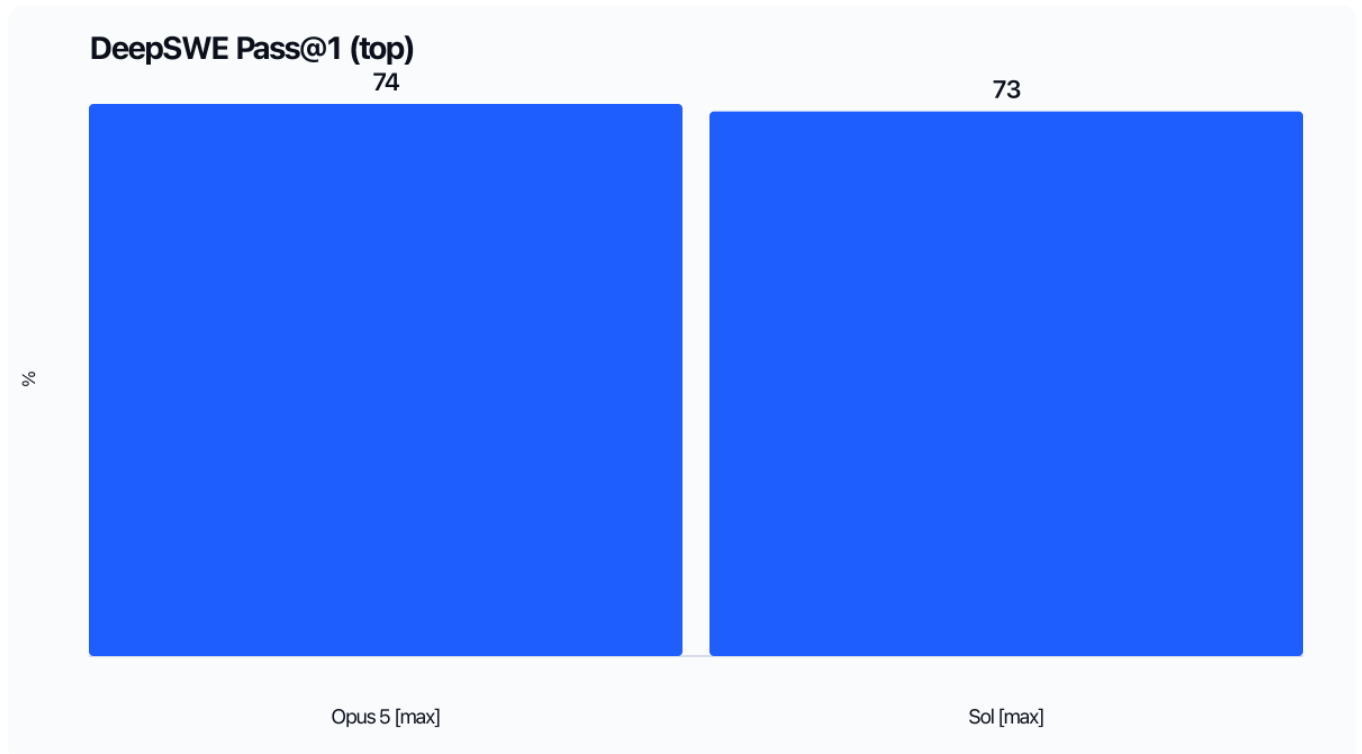
SWE-bench-style suites can saturate and leak when tasks are mined from public PRs. DeepSWE is Datacurve's long-horizon software-engineering agent benchmark built to reduce that contamination.

DeepSWE v1.1 grades committed patches in isolated verifier containers. The suite has **113 tasks** across **91 repos** and five languages; tasks are written from scratch with hand-written behavior verifiers, and reference solutions are not merged upstream. On the leaderboard (mini-swe-agent harness): Claude Opus 5 **[max]** at **74% ± 4%**, about



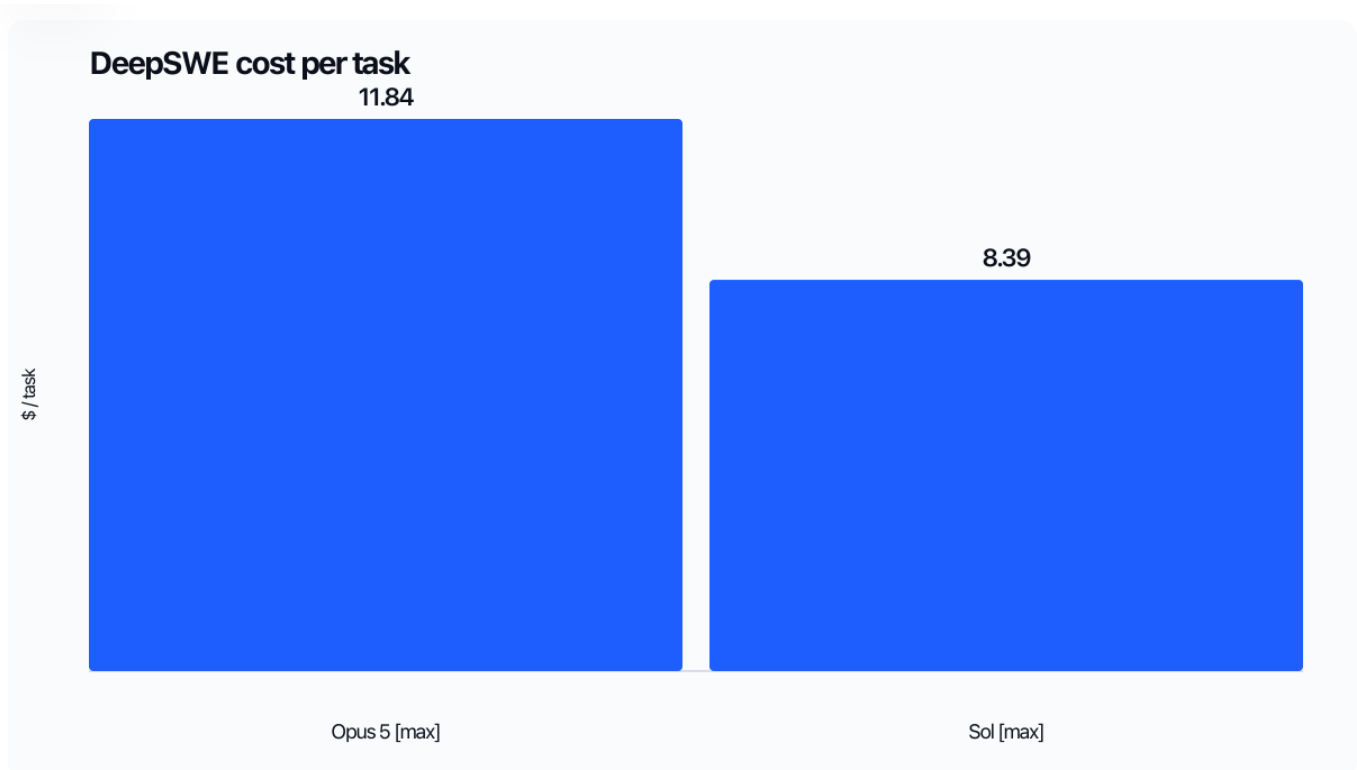
\$11.84/task, ~99 steps; GPT-5.6 Sol **[max]** at **73% $\pm$ 3%**, about **\$8.39/task**, ~61 steps. Confidence intervals overlap at the top.

This is not an ARC-AGI story. DeepSWE’s published contrast is with SWE-bench-style mined-task contamination — abstract reasoning puzzles versus multi-repo engineering are different domains. Do not import ARC scores into this section or the reverse.



DeepSWE Pass@1 (mini-swe-agent): Opus 5 max ~74% $\pm$ 4%, Sol max ~73% $\pm$ 3% — confidence intervals overlap at the top. See Sources below. [Source](#)





Cost vs Pass@1: Opus 5 max ~\$11.84/task, Sol max ~\$8.39/task. Cost-per-task is part of the scorecard. See Sources below. [Source](#)

Why it matters: long-horizon coding leaderboards are splitting into "saturated/mined" versus "contamination-aware," and cost-per-task is part of the scorecard now.

113 contamination-free tasks; Pass@1 with CI bands.

Opus 5 max ~74% ± 4% / \$11.84; Sol max ~73% ± 3% / \$8.39.

Contrast vs SWE-bench family — not vs ARC-AGI.

9. MCP 2026-07-28: stateless core (plus Claude's product rollout)

Older MCP assumed more sessionful, bidirectional behavior. A stateless request/response core lets any request hit any instance behind an ordinary load balancer — the same scaling model as normal HTTP APIs.

The **2026-07-28** spec ships that core. **`initialize` / `initialized` are retired; `Mcp-Session-Id` is removed.** Each request carries protocol version, client identity, and capabilities in `_meta`. Optional `server/discover` can fetch capabilities up front but is not required. App state still works via **explicit tool-minted handles** passed as arguments — not hidden transport sessions.



Multi Round-Trip Requests (MRTR) replace held-open server→client streams for elicitation/sampling/roots-style flows: the server can return `input_required`, and the client retries with `inputResponses`. Streamable HTTP **must** include **Mcp-Method** and **Mcp-Name** headers so gateways can route without parsing bodies. Extensions and auth updates (tasks extension, CIMD replacing DCR over time, enterprise-managed authorization) sit alongside a formal deprecation window for older features.

Secondary angle: Anthropic's Claude blog describes bringing MCP 2026-07-28 support across Claude products — connectors, MCP Apps, enterprise-managed auth, observability, tunnels — as product rollout on the same release, not a second protocol.

MCP 2026-07-28 — stateless core



Claude product rollout = same spec

MCP 2026-07-28: sessionful initialize / Mcp-Session-Id → stateless per-request _meta, explicit handles, and MRTR + Mcp-Method / Mcp-Name. Claude product rollout = same spec. See Sources below. [Source](#)

Why it matters: agent tool servers can scale like ordinary HTTP. Sticky-session assumptions become migration work. Claude aligning connectors to the same date is distribution, not a separate story.

Stateless core; initialize/session gone; explicit handles for state.

MRTR for multi-step input; Mcp-Method / Mcp-Name required on streamable HTTP.

Claude rollout = product support of the same spec.

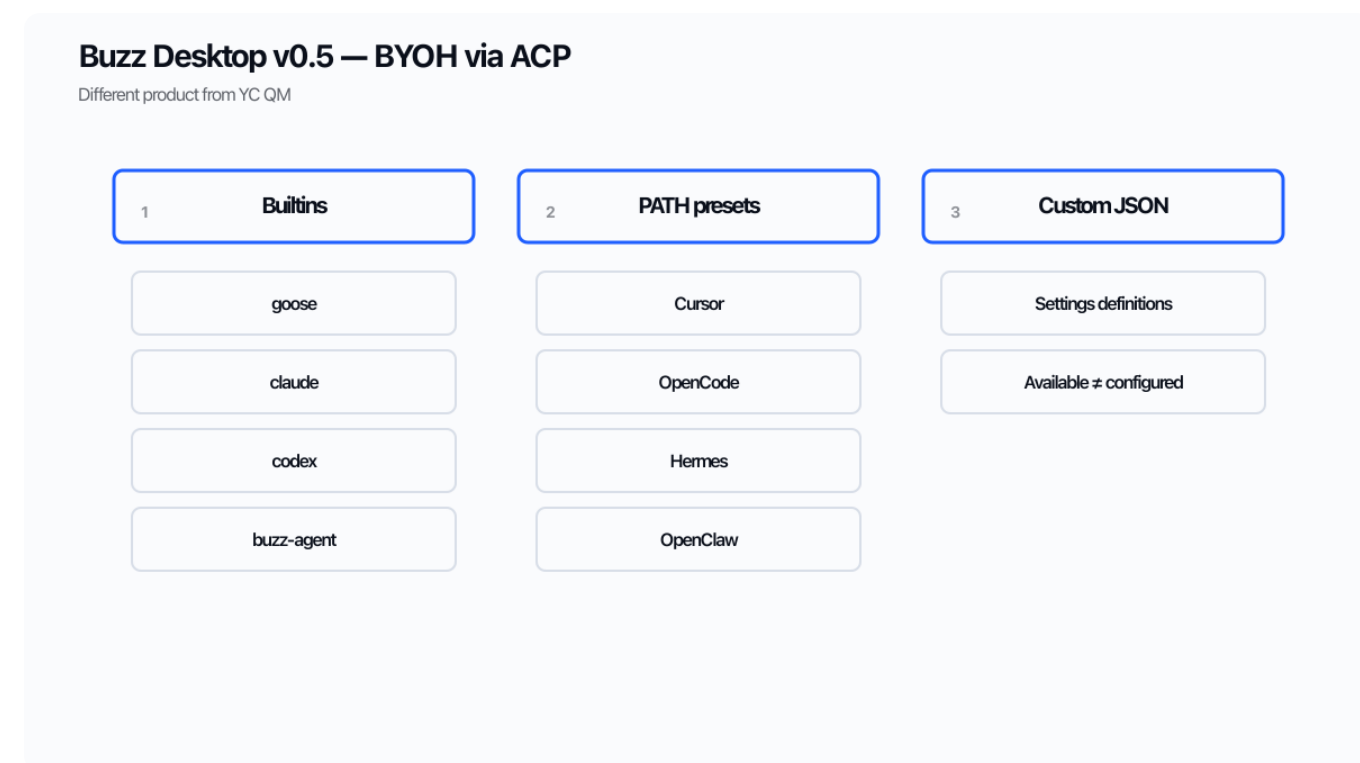


10. Buzz Desktop v0.5.0: Bring Your Own Harness via ACP

Desktop agent hosts used to ship one or two built-in runtimes. Bring Your Own Harness (BYOH) means the host speaks a shared Agent Client Protocol (ACP) seam and lets you plug in whichever coding agent is on your PATH.

Block's Buzz Desktop **v0.5.0** (28 July 2026) headlines BYOH via ACP (PR #2773). Runtime model: builtins (goose, claude, codex, buzz-agent); PATH-probed presets including **Cursor**, **OpenCode**, **Hermes**, and **OpenClaw** (plus others); and custom JSON definitions in Settings. Preset "Available" means the binary was found — not that auth and gateway setup are done (OpenClaw's ACP path is Gateway-backed).

Do not conflate Buzz with YC QM. Both name overlapping harnesses in the ACP ecosystem; Buzz is a desktop/community agent surface, QM is an org multiplayer Slack/web harness. Same wave, different products.



Buzz Desktop v0.5 BYOH via ACP: builtins (goose/claude/codex/buzz-agent), PATH presets (Cursor/OpenCode/Hermes/OpenClaw), custom JSON — Available ≠ configured. Different product from YC QM. See Sources below. [Source](#)

Why it matters: pluggable harnesses are becoming table stakes for agent workspaces — hosts compete on UX and ops, not on locking one model loop.

 BYOH via ACP; presets include Cursor / OpenCode / Hermes / OpenClaw.

Available ≠ fully configured.

Different product from QM despite shared harness names.

11. Anthropic on open weights: no category ban — three pillars

“Open weights” means releasing model parameters so others can run and fine-tune locally. Policy fights mix national-security export controls, IP/distillation accusations, and safety testing. Anthropic’s post is advocacy, not law.

On 27 July 2026 (with a 28 July edit note), Dario Amodei published Anthropic’s position. Explicit line: Anthropic has **never advocated a category ban** on open-weights models; non-dangerous open weights are framed as a public good. The post answers reports of possible U.S. restrictions on Chinese open-weights use and an industry open letter supporting open weights.

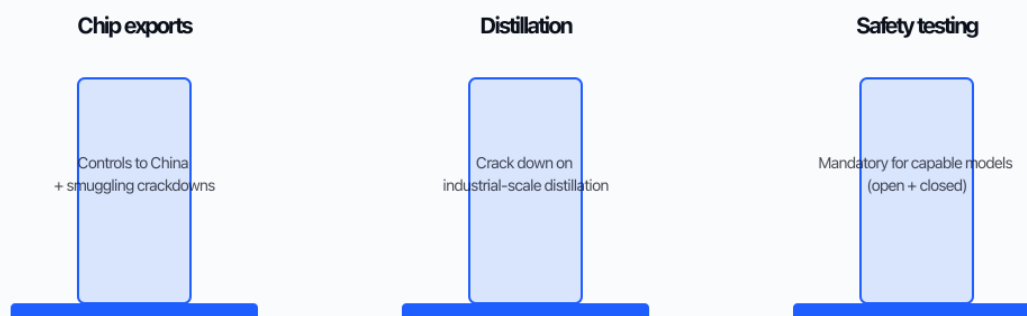
The three policy pillars: (1) **chip and chipmaking export controls** to China, plus crackdowns on smuggling/workarounds; (2) crack down on **industrial-scale distillation**; (3) **mandatory safety testing** for sufficiently capable models — **open and closed** — with weaker startup/academic models exempt. Anthropic agrees with parts of the industry letter on access and competition, and disagrees that open weights necessarily help defenders (especially bio offense/defense asymmetry). Cross-link last week’s White House distillation allegations around Kimi K3 as political context — do not re-litigate them here.



Anthropic open-weights position

Policy advocacy — not enacted law

No category ban — non-dangerous open weights as public good



Anthropic open-weights stance: no category ban; pillars = chip exports, industrial distillation crackdown, mandatory safety testing (open + closed). Policy advocacy — not enacted law. See Sources below. [Source](#)

Why it matters: clarifies Anthropic's public stance in a live China/open-weights fight — targeted compute, distillation, and testing, not "ban open weights."

No category ban; open non-dangerous weights framed as public good.

Pillars: chip exports, industrial distillation, mandatory safety tests (open + closed).

Policy advocacy — not enacted statute.

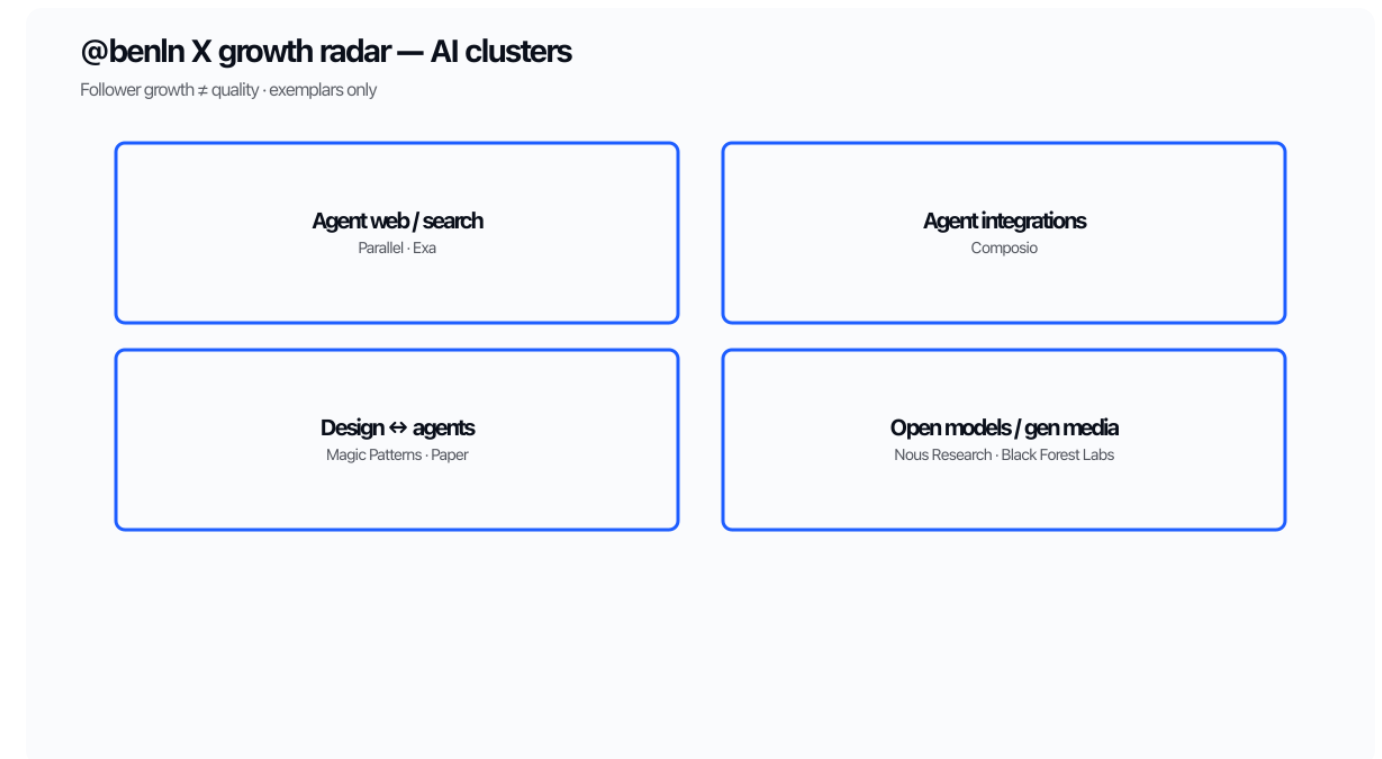
12. X growth radar: follower growth $\neq$ quality

Rankings of "fastest-growing startups on X by follower growth" measure attention under a platform's ranking rules. They are useful as a noisy radar for which AI themes are winning distribution — and useless as diligence.

@benln posted a list of the fastest-growing startups by X follower growth over about 90 days (full list image-only in mirrors; length in the high-20s to ~35 is plausible). Date caveat: some mirrors timestamp the post around **7 June**, while week-of briefs sometimes say ~26 July — treat the calendar day as uncertain and the metric as the point. We are **not** republishing the full ranked list.



AI-relevant clusters visible around the thread: agent web/search (Parallel, Exa), agent integrations (Composio), design↔agents (Magic Patterns, Paper), and open models / generative media (Nous Research, Black Forest Labs). Exemplars only — not an endorsement ranking:



@benIn X growth radar — AI clusters only (exemplars): agent web/search, integrations, design↔agents, open models/gen media. Follower growth ≠ quality. See Sources below. [Source](#)

Parallel — web search/research APIs built for agents.

Exa — search/crawl/research API for agent workflows.

Composio — large app-integration and auth layer for tool calls.

Magic Patterns — prompt → UI prototypes (self-ID as #28 on the list).

Paper — design canvas tied to agents/code/repos.

Nous Research — open-source LLMs and training infra.

Black Forest Labs — frontier image/video models (FLUX family).

Why it matters: this week's agent-memory, OSS, GPU, and web-agent themes show up in social attention. Growth can be ads or virality. Use it as a map of what people are watching — never as a quality or revenue score.

M ⚡ = follower growth on X, not product quality.

Date uncertain (mirror may say Jun 7) — do not overfit the week label.

Pattern clusters + a few exemplars; no dump of ~35 names.

Frequently Asked Questions

Is YC QM a finished product I should run in production?

No. Y Combinator itself calls QM an experiment that is early and has bugs. It is MIT open source with Slack and web surfaces, pluggable harnesses, and Strict/Auto/Dangerous security postures — useful to study and dogfood, not a polished SaaS guarantee.

Where can I use Google Vids personal avatars?

Per Google Workspace Updates and Help: 18+, English only, and not available in the EEA, Switzerland, or the United Kingdom. Eligible plans include Google AI Pro/Ultra and listed Workspace SKUs. Every generated clip carries SynthID. Rapid Release started July 16, 2026; Scheduled Release August 5, 2026.

Did SpaceX already close the Cursor deal?

Not as confirmed closed in the sources we cite. On June 16, 2026, SpaceX and Anysphere signed a merger agreement at an implied ~\$60B all-stock equity value, with expected close in Q3 2026 subject to approvals. Signed ≠ closed.

Did Mem0 cut total Claude Code tokens by 97%?

No. Mem0's first-party demo reports memory footprint falling from ~13,700 tokens to ~445 (~97%). Total session tokens in the same table went from about 75k to 67k — a much smaller drop. Treat it as a vendor experiment, not an independent audit.

Is GPT-5.6 Sol's 38.3% ARC-AGI-3 score the official leaderboard number?

No. OpenAI reports ~13.3% on the official public harness versus ~38.3% with retained reasoning plus compaction on the public set. Official ARC Prize semi-private for Sol Max is about 7.78%. Do not mix those scores. François Chollet has said general-purpose API settings are acceptable if disclosed — still not the same as the official harness.



Has Superlogical shipped a product?

Not as of the July 29, 2026 announcement window we cite. The site is a beta waitlist for a terminal multiplexer built on libghostty. The broader 'multiplexer for all work' thesis is stated; the product is not out yet.

What changed in MCP 2026-07-28?

The core became a stateless request/response model: initialize/initialized and Mcp-Session-Id are gone. Multi Round-Trip Requests (MRTR) replace held-open server→client streams for elicitation-style flows. Streamable HTTP must send Mcp-Method and Mcp-Name headers. Claude is rolling support across products as a secondary product angle to the same spec.

Is Buzz Desktop the same as YC QM?

No. Both sit in the pluggable-harness wave and name overlapping runtimes (Cursor, OpenCode, Hermes, OpenClaw), but Buzz Desktop (Block) is a desktop/community agent host with BYOH via ACP, while QM is YC's org-scoped multiplayer harness for Slack and web. Different products.

Sources

This is original synthesis. Underlying facts and charts are drawn from the reporting and vendor posts below.

[↗ YC QM homepage](#)

[↗ GitHub — yc-software/qm](#)

[↗ Google Blog — Gemini Omni and personal avatars in Vids](#)

[↗ Workspace Updates — personal avatars with Gemini Omni in Vids](#)

[↗ TechCrunch — Google Vids AI video / avatars](#)

[↗ TechCrunch — Cognition acquires Windsurf remainder](#)

[↗ CNBC — Google Windsurf reverse-acquire](#)

[↗ ⚡ Cognition — Windsurf blog](#)

[🔗 GitHub — Roo-Code \(archived\)](#)

[🔗 Roomote](#)

[🔗 SEC — SpaceX / Anysphere merger 8-K](#)

[🔗 CNBC — SpaceX to buy Cursor parent Anysphere for \\$60B](#)

[🔗 Continue.dev](#)

[🔗 The New Stack — Cursor acquires Continue](#)

[🔗 Anaconda — acquires Kilo Code](#)

[🔗 The Information — OpenAI hires Cline staffers](#)

[🔗 Kilo blog — Cline clarification note](#)

[🔗 Mem0 — How Mem0 cut Claude Code's memory footprint by 97%](#)

[🔗 Mem0 docs — Claude Code integration](#)

[🔗 Microsoft — FY26 Q4 earnings press release](#)

[🔗 SEC — Microsoft EX-99.1 FY26 Q4](#)

[🔗 OpenAI — How two settings tripled our ARC-AGI-3 scores](#)

[🔗 ARC Prize — GPT-5.6 results](#)

[🔗 The Decoder — OpenAI ARC-AGI-3 harness debate](#)

[🔗 Superlogical](#)

[🔗 Mitchell Hashimoto — Superlogical](#)

[🔗 DeepSWE leaderboard](#)

[🔗 DeepSWE v1.1 blog](#)

[🔗 MCP Blog — Spec 2026-07-28](#)

[🔗 Claude — Bringing MCP 2026-07-28 to Claude](#)

[🔗 GitHub — block/buzz v0.5.0](#)

[🔗 GitHub — Buzz BYOH PR #2773](#)

[🔗 Anthropic — Our position on open-weights models](#)

[🔗 Ben Lang X post mirror — fastest-growing startups by follower growth](#)

Sections on Mem0 report a first-party vendor experiment, not an independent audit. Microsoft figures are educational corporate results — not investment advice. ARC-AGI-3 scores depend on harness and must not be mixed. Coding-tool consolidation claims are fact-checked against primary reports; social posts are hooks, not sources of truth. Anthropic's open-weights piece is policy advocacy, not enacted law. X follower-growth rankings measure attention, not quality. This digest is educational synthesis of public reporting and vendor posts — not legal, security, or financial advice. Always verify against primary sources.

